Local Air Quality and Economic Inequality in Finland

Nina Hiltunen

School of Business

Thesis submitted for examination for the degree of Master of Science in Economics and Business Administration.
Espoo 27.4.2020

Supervisor

Asst. Prof. Tuukka Saarimaa

Advisor

Prof. Matti Liski

Aalto University
School of Business

Aalto University
School of Business

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi
Abstract of the master's thesis

Author Nina Hiltunen

Title Local Air Quality and Economic Inequality in Finland

Degree programme Economics and Business Administration

Major Economics **Code of major** BIZ24

Supervisor Asst. Prof. Tuukka Saarimaa

Advisor Prof. Matti Liski

Date 27.4.2020 **Number of pages** 41 **Language** English

Abstract

This master's thesis calculates the distribution of an adjusted income in Finland for 2015. Adjusted income is calculated by deducting damages from exposure to fine particles (PM2.5) from average market income. When including the damages into annual income measurement, the income distribution becomes less evenly spread. This suggests that the damages are not distributed equally within the income quintiles. Contrary the study made by Muller et al. (2018) in United States, the PM2.5 levels in Finland are on average higher the more affluent the area is. However, the lower income quintiles experience higher damages. The main reason for this lies within the age of population in these quintiles: lower income quintiles have higher shares of older population, who inherently are more vulnerable to air pollution.

This leads to a change in the share of income in income quintiles: the top quintile relatively gains 0.7 percentage points while the bottom quintile loses 0.5 percentage points. Only the middle quintile is not affected by the inclusion of the air pollution externality.

Damage per market income values show how many percent is the average damage relative to average income for the population in a grid belonging to that area. Interestingly enough the maximum share is not the highest in Southern Finland, which has the highest maximum for PM2.5 levels. The higher the damage/market income ratio is only for a certain group of population, the higher will be the Gini coefficient for adjusted income.

The monetary benefits from fine particle reduction should be taken into consideration when policies regarding air quality and health of the population are considered. When the PM2.5 levels are reduced by 30 percent, total savings for the year 2015 would lead to 1.1 billion euros and the Gini coefficient for adjusted income would drop from 0.155 to 0.150. However, as the population of Finland is getting older, the savings would be even bigger in the future. From the decreased Gini coefficient and the savings in damages we can separate two channels: income effect and redistribution effect. So, environmental policy aiming to decrease the level PM2.5 can increase the level of economic welfare and change the distribution of said welfare more equal. According to the law of diminishing marginal utility, total welfare would be increased.

Keywords Economic Inequality, Air Quality, Gini Coefficient, Value of a Statistical Life

Contents

1 Introduction

Air pollution has been recognised as one of the leading environmental health risks globally. Particulates are the most harmful form of air pollution due to their ability to penetrate into the lungs and blood stream, causing possible increased symptoms and in the worst case, a premature death. According to a research made by National Institute for Health and Welfare (THL), Finnish Environmental Institute (SYKE) and Finnish Metrological Institute (IL), air pollution was associated with 2,000 premature deaths (1,400–2,600 95% CI) for year 2015 in Finland. Fine particles (or PM2.5) were the main contributor (74%) to the disease burden and most of the burden could be attributed to mortality with 1,600 premature deaths and 32,900 years of life lost. [Lehtomäki et al., 2018] PM2.5 means atmospheric particulate matter (PM) that has a diameter of less than 2.5 μm (micrometres), which is roughly 3 percent of the diameter of a human hair.

To put some magnitude to the effect, THL estimated that in 2012 tobacco smoking contributed to around 4,200-4,500 deaths [Vähänen, 2015]. However, smoking can't be compared to air pollution exposure, as the latter is treated as an externality, while the adverse effects from smoking are self-inflicted. Secondhand smoking, on the contrary, is also an externality and is included in air pollution. Another study made by THL notes that PM2.5 causes around 4,500 DALYs/year (disability-adjusted life years) in 1 million people while secondhand smoke causes 600 DALYs/year [Hänninen et al., 2014]. One DALY can be thought of as one lost year of healthy life, free of disease or disability [World Health Organization, 2020].

Epidemiologic research has estimated a quantitative relationship between exposure to high concentrations of fine particulate matter (PM2.5) with increased mortality or morbidity. Increased mortality means higher incidence of deaths in an exposed population, while morbidity refers to having a disease or the amount of disease in a population. From these have been concluded an exposure–response relationship that describes the increased negative effects with increasing air pollution concentration level. [Lehtomäki et al., 2018] World Health Organization (WHO) has considered air pollution as a health risk since 1958, but the lack of data and related methodologies prevented development of specific recommendations at that time [World Health Organization, 1958]. The first air quality guidelines for Europe were published in 1987. In 2000, PM2.5 was given an exposure–response relationship without a threshold [World Health Organization, 1987] [World Health Organization, 2000]. WHO set a guideline level for PM2.5 to an annual mean of 10 $\mu g/m^3$ (micrograms per cubic metre) for the first time in 2005[World Health Organization, 2006]. This guideline is based on evidence of increases in heart and lung disease mortality, with more than 95 percent confidence, due to long-term exposure to PM2.5 concentrations higher than 10 $\mu g/m^3$. Long-term exposure is defined as exposure lasting a year or more. [Hoek et al., 2013] Since then there has been growing evidence of adverse health effects at lower concentrations [Crouse et al., 2012]. There is no identified threshold below which no damage to health is observed. Hence, the WHO 2005 guideline aimed to achieve the lowest concentrations of PM2.5 possible. [World Health Organization, 2018]

Quantifying the adverse health effects of fine particles helps policymakers to estimate expected health improvements in the population if PM2.5 were to be reduced. The Air Pollution Damage Cost Model for Finland (IHKU) is the first integrated tool for estimating the economic costs of Finnish air pollution emissions [Savolahti et al., 2018]. The IHKU model includes premature death, short- and long-term illnesses and the decrease of work performance in the cost calculation.

Premature death forms the greatest part of the total cost, so the estimations vary based on the assumed cost of premature death. Two metrics often used are the Value of Statistical Life (VSL) and the Value of a Life Year (VOLY). VSL uses the estimate of willingness to accept extra income for a small increase in mortality risk. VOLY is derived from VSL by dividing it with the expected value of remaining life years.

Most of the research in Finland has concentrated on the health damage costs of Finnish air pollutants, but not how it could possibly affect economic inequality. The most commonly used measurement of economic inequality is income distribution. Recent research in the United States has supported two conclusions; with access to more precise data sets, there is now an accurate sense of income distribution and that the distribution has become more unequal over time [Muller et al., 2018]. Muller et al. (2018) challenge these conclusions by arguing that the distribution is actually much more unequal than the studies have suggested but it might be also becoming more equal over time. They point out that the improved measurement of income doesn't only depend on better records of market income, but also on more comprehensive definitions of income. This thought is based on the suggestion made by Nordhaus and Tobin already in the 1970s. [Muller et al., 2018] Nordhaus and Tobin argued that the official focus on output – and therefore income– of goods and services produced for, and sold in, organized markets is too restrictive [Nordhaus and Tobin, 1972]. A decade later, Blinder (1980) argued that if we measure income more comprehensively and over periods longer than a year, a clearer trend toward equality would emerge [Blinder, 1980].

Muller et al. (2018) calculate the distribution of an adjusted measure of income in the United States from 2011 to 2014. The adjusted measure deducts damages due to exposure to air pollution from reported market income. This specification is based on the principle of adding environmental externality into measures of national income that has been found in a longstanding literature. They are the first to associate individual and household damage with market income and to construct the distribution of said measure of adjusted income. The results of the study showed that inclusion of air pollution damage acts like a regressive tax: with air pollution, the bottom 20 percent of households lose roughly 10 percentage points of the share of income, while the top 20 percent of households gain 10 percentage points in 2014. The Gini coefficient for this measure of adjusted income is 0.682 in 2011, as compared to 0.482 for the original market income. [Muller et al., 2018]

The same model used in the previously mentioned study is going to be applied in a Finnish context with slight modifications. These modifications are due to some data restrictions in the available Finnish data. The Gini coefficient increases from 0.141 to 0.155 when adjusted income is used instead of the original market income.

The Gini coefficient is almost half of the official Gini coefficient of Finland. This is due to the degree of data precision. However, more relevant information is how does the Gini coefficient change from the original when externality from exposure to PM2.5 is included in the measurement. The Gini coefficient for externality is 0.388, which means that the damages are concentrated on a subset of grids. By inspecting the population distribution in income quintiles, the results show that the lower quintiles have larger share of older population than the higher quintiles. Even though the pollution levels are lowest in the lower income quintiles, the inherent vulnerability to air pollution rises the damages high in them.

Damage per market income values show how many percent is the average damage relative to average income for the population in a grid belonging to that area. Interestingly enough the maximum share is not the highest in Southern Finland, which has the highest maximum for PM2.5 levels. The higher the damage/market income ratio is only for a certain group of population, the higher will be the Gini coefficient for adjusted income.

Another thing to consider is the possible effect of PM2.5 level reduction on mortality risk and it's relation to income distribution. The European Union has set a National Emission Ceilings Directive that sets national emission reduction commitments for Member States for five air pollutants, including PM2.5, for year 2030. In 2019, the Finnish Ministry of the Environment updated the emission reduction obligations set by the directive which includes reducing PM2.5 with 30 percent till the end of year 2029. Therefore, with a reduction of 30 percent in PM2.5 levels in all grids, the Gini coefficient for the adjusted income will decrease to 0.150. The total savings would be 1.1 billion euros which would even increase in the future, as the population in Finland is getting older.

Lastly, economic inequality can show up in housing prices. It is interesting to see, whether there is any kind of housing segregation between population from different socio-economic backgrounds. This would have significance from the viewpoint of economic equality, if poorer population ended up living in more polluted areas, thus incurring more damage from exposure. In Helsinki, cheaper accommodations seem to reside on more polluted areas on average. Based on the average housing prices and average PM2.5 levels, it is a possibility that income-constrained population can't choose to live in a less polluted area as freely as a richer population could, resulting in larger monetary damages from exposure.

2 Measuring Income

Muller et al. (2018)'s adjusted income measure deducts monetary costs attributable to air pollution exposure from reported household income. This specification is based on the principle of adding environmental externality into measures of national income that has been found in a longstanding literature. [Muller et al., 2018]

Nordhaus and Tobin argued that the official focus on output – and thus income– of goods and services produced for, and sold in, organized markets is too restrictive [Nordhaus and Tobin, 1972]. Later on, Nordhaus introduced the concept of augmented national accounts during the 1990s. Augmented accounts is defined as an integrated set of accounts for both market and non-market economic activity. These accounts should measure income and output in a manner that best correspond to net economic welfare. In other words, they should include both market and non-market activities. Areas of importance include natural resources, unpaid work, leisure time, investment in education and health, and the environment among other things. [Nordhaus, 2006] The U.S. Bureau of Economic Analysis developed the Integrated Economic and Environmental Satellite Accounts because of the growing importance of environmental accounting globally. Environmental accounting provides valuable information about the interaction between the environment and the economy. It provides information on the implications of different regulations, taxes, and consumption patterns among other things. [National Academy of Sciences, National Research Council, 1999]

A decade later Blinder (1980) argued that estimating the trend in economic equality depends on what is included in the measurement of income and length of the accounting period. To address the problem in the method of income measurement, he chooses as an example the post-World War II period in the U.S. when income equality was about the same in 1977 as it was in 1974, according to the official data. However, he reminds us that the American population experienced substantial demographic changes during this thirty-year period. Given the way income distribution data is compiled, there would have been substantial trend towards greater inequality with these demographic shifts if other factors had not intervened. Government transfer programs played a major role in this. For example, when higher living standards and/or more generous public transfer programs enable grandparents to move into apartment of their own, a new economic unit is formed with rather low income. This brings down the average level of income and raises inequality. [Blinder, 1980]

Blinder (1980) also criticises the default accounting period. Why are we interested in distributions of annual income instead of incomes measured over some alternative accounting period? However, a day or a week would be a too short period to generate meaningful data on income inequality as most people have weeks of zero income without being poor. A year can also be a too short accounting period to place many people meaningfully within the income distribution. Blinder uses as an example investment in human capital, since it typically leads to rising age-earnings profiles, people appearing to be poor during certain years, but over a lifetime being well off. This is why issues like defining and measuring income are important when measuring economy's performance. [Blinder, 1980]

Being interested in the level and distribution of income means that we believe these two values are two approximate indicators of economic welfare. Higher or more equally distributed income means that society is better off. Most use the current income to summarize the whole opportunity set. But we know that two individuals with identical opportunities will have different incomes if they have different preferences. Often ill health requires more current income to achieve any given level of satisfaction. [Blinder, 1980] These points can be applied also to environmental externalities, as the exposure to pollution can lead to illnesses over time and therefore incur monetary damages.

Despite the wide addressment of environmental externalities in earlier works, Muller et al. (2018) are the first to associate individual and household damage with market income and to construct the distribution of said measure of adjusted income. With the available data of PM2.5 concentrations and socio-economic data within Finland, it is possible to try apply the methods in the Finnish context. While the socio-economic data is not on the most precise level possible due to data privacy reasons, the PM2.5 concentration data is estimated with the Finnish Regional Emission Scenario model, point estimates based on a comprehensive set of emission monitors.

3 The Adverse Health Effects from PM2.5 Exposure

Cohen et al. (2017) estimate the population-weighted mean concentrations of PM2.5 globally by aggregating grid-level exposure concentrations to national-level population-weighted means using the corresponding grid cell population value. Their results show Finland being one of countries with lowest exposure to PM2.5, as expected (less than 8 $\mu g/m^3$). In comparison, the same measurement in China is 58.4 $\mu g/m^3$ and the highest one being in Qatar 107.3 $\mu g/m^3$. [Cohen et al., 2017] However, as mentioned earlier, there are adverse health effects even with low PM2.5 levels.

Health impacts of air pollution increase with age, especially above 30 years of age (Figure 1a). After infancy, mortality is rare at younger age groups (5–30 years), while for other age groups it is the dominant endpoint (Figure 1b). [Lehtomäki et al., 2018] The mortality rate at some age x refers to the probability of a person living until the age of x to die during that year of age [Statistic Finland, 2020c]. Death is considered to be premature when a person dies before their life expectancy.

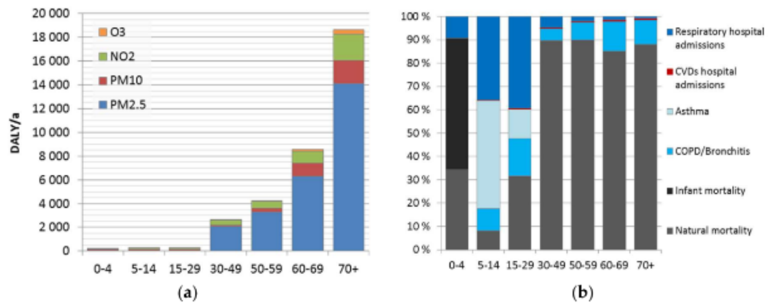

Figure 1: a) Disease burden of three air pollutants in seven age groups in Finland in 2015. b) Health outcomes caused by ambient air pollution in age groups. [Lehtomäki et al., 2018]

Long-term, over years of exposure is especially harmful because not only can it affect sudden illness, it can also speed up the development of an illness overtime or even cause it. This leads to long-term effects being substantially higher than short term effects. The long-term adverse effects can be estimated epidemiologically by comparing the incidence of illnesses or the probability of falling ill between areas with different air pollution concentrations. Some of the illnesses appear after years or decades, but some can be observed with a lag of a couple of years. [Savolahti et al., 2018]

Particle pollution is generally considered to be more harmful than gas-like emissions. The adverse effects of fine particles are more well-known than ultra small particles or coarse particles (such as street dust). What adds to the harmfulness of fine particles is that they spread to a large area. In addition, fine particles get into buildings and penetrate the lungs. Due to these reasons the impact analyses of air pollution tend to concentrate on fine particles. [Savolahti et al., 2018]

A study by Laden et al., suggests that mortality effects may be partially reversible over a time period as short as a year. This then suggests that ambient PM2.5 is likely associated with exacerbation of existing disease. In addition, there also appears to be a second independent effect that can be described as development of chronic disease. [Laden et al., 2006] World Health Organization's Health risks of air pollution in Europe (HRAPIE) -project experts recommend to express the benefits of reduced exposure to PM2.5 is in terms of life-years gained across the population as a whole [World Health Organization, 2013].

4 Data

The socio-economic and demographic data from Statistics Finland, provided by the city of Helsinki, is on a grid-level of 250 m x 250 m, based on ETRS89-TM35FIN coordinates. It includes the number of inhabitants in every age cohort for year 2014, the amount of adult population for year 2013 and the average disposable income of inhabitant for year 2013. Disposable money income includes monetary income items and benefits in kind connected to employment relationships, like wages and salaries, entrepreneurial income and property income [Statistic Finland, 2020a]. Property income includes rental, interest and dividend income derived from registers, taxable capital gains and pensions based on private insurance, and other property income derived from taxation data [Statistic Finland, 2020d].

The population data for year 2014 will be used, as it contains more precise information about the structure of the age cohorts, and will be calculating the amount of adult population per grid by summing age cohorts older than 17 years old. Even though the years are not the same, the assumption is that they are relatively similar in year 2015. The total number of observations is 627,936 but due to information privacy, any grid including less than 10 inhabitants will have their income information hidden. Thus, the usable amount of observations is reduced to 53,871. This means that out of adult population of 4.3 million, income and demographic data for 3.6 million adults will be used. Supplementary table A. 5 has the exact numbers for both years 2013 and 2014 and A. 6 the regional population for year 2014. In the Robustness -subsection the change in results is tested when grids with higher amounts of population are dropped. Supplementary table A. 1 summarizes the population size by age cohorts in different grid-cells.

The data for PM2.5 concentrations in Finland for year 2015 is provided by Finnish Environmental Institute. The concentrations are estimated with the Finnish Regional Emission Scenario (FRES) model with 250 m spatial resolution for areal sources with ETRS89-TM35FIN coordinates and an comprehensive list of point emissions (n= 581) in Finland. [Karvosenoja, 2008] Table 1 summarizes the average disposable income and the levels of PM2.5 concentrations. Only the areas with population and income data are used, as based on those concentrations the damages are calculated.

Table 1: Disposable Income and PM2.5 Levels

	Income	PM2.5
Mean	24,049	5.238
Median	22,737	5.238
SD	10,495	1.348
Min.	340	0
Max.	1,017,969	12.485
Obs.	53,871	53,871
Year	2013	2015

The highest level of PM2.5 exceeds the annual recommendation for fine particle level 10 $\mu g/m^3$ made by WHO. The biggest sources of fine particles in Finland come from wood burning in households and traffic. Figure 2 illustrates the annual amount of PM2.5 pollution in kilograms. [Savolahti et al., 2018] Over half of PM2.5 levels are due to Finnish regional air pollution dispersion and far distance dispersion.

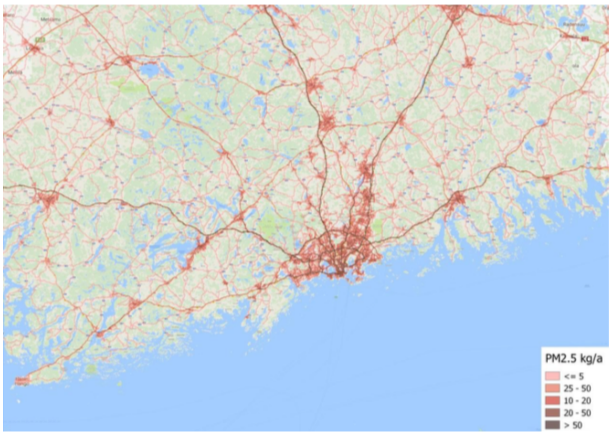

Figure 2: PM2.5 Pollution Caused by Traffic [Savolahti et al., 2018]

Figure 3 is based on the FRES-model that has included the effect of PM2.5 dispersion on regional emissions [Soimakallio et al., 2017]. As we can see from the picture, there is a high concentration of PM2.5 pollution around Saint Petersburg that possibly affects the pollution levels in Finland.

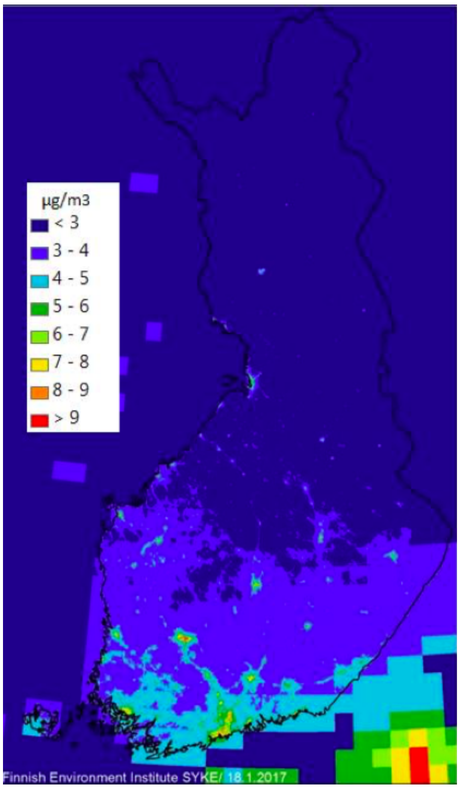

Figure 3: Modelled PM2.5 Concentration Levels [Soimakallio et al., 2017]

The two following histograms 4 and 5 portray the distribution of PM2.5 levels in both uninhabited and inhabited areas. In both cases, the levels are concentrated around 3-5 $\mu g/m^3$ with inhabited areas having larger share of higher PM2.5 levels as expected. The red dashed line at 10 $\mu g/m^3$ presents the WHO annual PM2.5 concentration level recommendation. Within both areas, there are some grids that exceeds this recommendation. In uninhabited areas, these grids locate in highways while the inhabited grids are nearby these polluted highways.

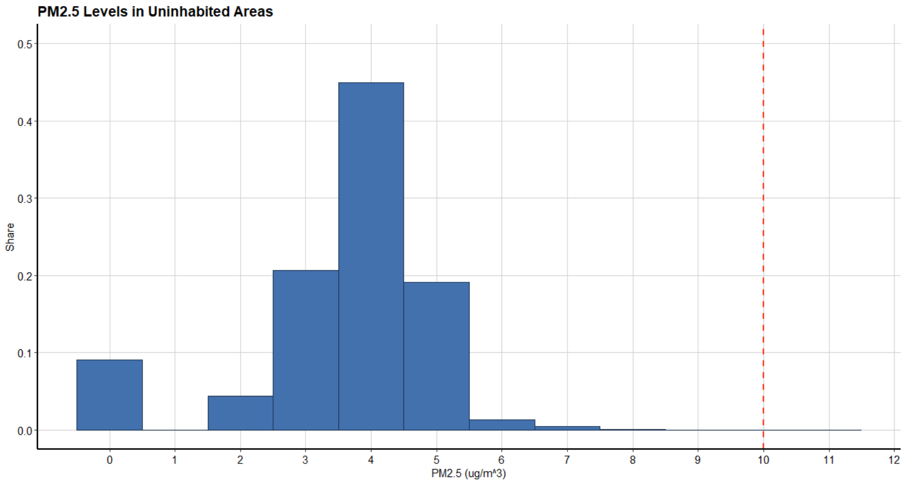

Figure 4: PM2.5 Levels in Uninhabited Areas

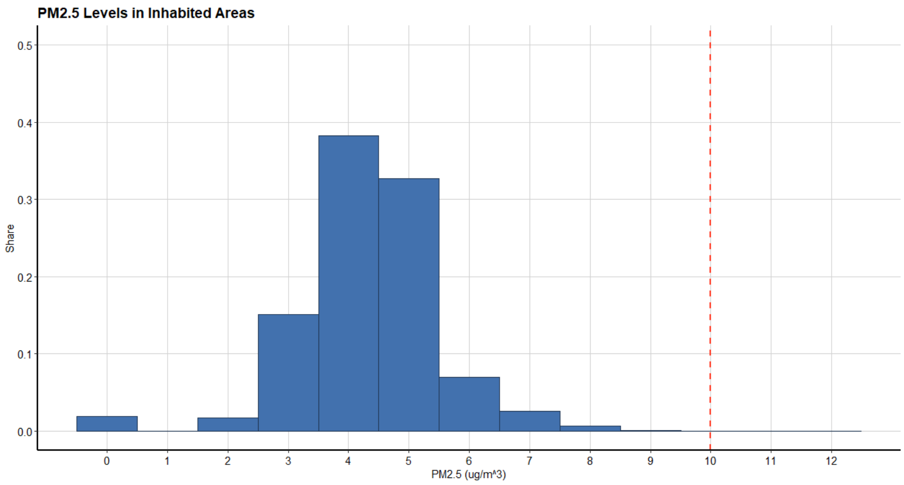

Figure 5: PM2.5 Levels in Inhabited Areas

The public baseline health statistics from Statistics Finland includes country-specific life expectancies and all-cause mortality rates for the year 2015, stratified by age and sex. The all-cause mortality data is not stratified by the reason of death, so it might include deaths from violent causes of death and injuries. The data is aggregated to age cohorts without separating the sexes because the health statistics are added to the earlier mentioned population statistics. Supplementary table A. 2 shows the calculated baseline mortality risks by age cohort.

HRAPIE experts recommend using all-cause analysis as primary choice over using specific causes of death. Usually the national data on all-cause mortality has a greater precision than the cause-specific data. This is due to the possibility of the latter being affected by misclassification of causes of death in mortality registration. Evidence supports that there is a relationship between cardiovascular, respiratory and lung cancer mortality and PM2.5 exposure, while many other causes of death (such as gastrointestinal diseases) are less likely affected by air pollution. However, the frequency of the causes of death associated with exposure was similar enough between cohorts included in meta-analysis and between countries in the EU where the impacts are to be estimated. Therefore it's justified to use all-cause mortality. [World Health Organization, 2013]

5 Monetizing the Damages

5.1 Metrics to Quantify the Health Impacts from Exposure

There are many ways to quantify the health impacts of a population due to exposure to air pollution. Often the results are reported in terms of the number of attributable deaths or the cases of disease, years of life lost (YLL) or disability-adjusted life years (DALYs). [World Health Organization, 2016] In this study, it's sensible to use the number of attributable deaths and more extensively years of life lost, as the data available is restricted to mortality rates and expected life years. In addition, previous literature has stated that most of the monetary costs come from premature deaths, not so much from disabilities caused from the exposure.

As premature death forms a greater part of the costs, the estimations vary based on the assumed cost of a premature death. Metrics describing the monetary damage of premature death are not affected by the person's contribution to society. It's based on a revealed preference of consumer behaviour as preferences of population can be revealed by their willingness to pay for an decrease in mortality risk.

Two different metrics are used for the estimation of value of life. The Value of a Statistical Life (VSL) uses the estimate of willingness to accept extra income for small increase in mortality risks. For example, if a person accepts 600 euros for an increase of 0.01 percent in mortality risk, the VSL would be 600 euros/0.0001 = 6,000,000 euros. This metric assumes that every life is equally valuable, not depending on how many life years a person is still expected to have.

The other one is the value of a life year (VOLY) that is often derived from VSL. The assumption is that the remaining years are all equally valuable. VOLY is calculated by dividing VSL with the expected value of remaining life years i.e.the YLL metric. This means that younger population that has more expected life years left, has higher value of life. There has been no international consensus which metric should be used. [Savolahti et al., 2018]

In this thesis, the VSL value of 2.65 million euros will be used as a default metric and value as it has been used in the previously mentioned Finnish study made by THL, IL and SYKE. It will be also tested how the Gini coefficient changes when different value and metric is used. The European Environmental Agency (EEA) recommends using values from their study NewExt, conducted in year 2000. The study was conducted in three different countries. The average VSL was 2 million euros, median VSL 0.98 million euros, average VOLY 120,000 euros and median VOLY 52,000 euros. These values were inflation-adjusted for Clean Air Act-programme which resulted in the values 2.65 million euros, 1.3 million euros, 160,0000 euros and 69,000 euros. Interestingly enough the European values are systematically lower than the American ones which can't be explained with the differences in standard of living. In 2016, United States Environmental Protection Agency (USEPA) recommended using 10.3 million dollars (base year 2013) that should be adjusted to price level and use the income elasticity of 0.7. On the contrary, EAA has not decided to adjust the values with the level of income. [Savolahti et al., 2018]

A concentration-response function represents the risk of air pollution exposure to health in a population. This function is typically based on Relative Risk (RR) estimates that are derived from epidemiological studies. The RR estimate describes the probability of an adverse health outcome, in this study the likelihood of a premature death, when a population is exposed to a higher level of air pollution relative to a population with a lower exposure level. [World Health Organization, 2016] Often RR is expressed as the proportional increase in the estimated health outcome associated with an increase in pollutant concentrations. [World Health Organization, 2016] The concentration-response coefficient varies from 2.14 per cent to 10.6 per cent for 10 μg increase in many previous literature. WHO has recommended the value of 6.2 per cent or 0.0062 for 1 $\mu g/m^3$ to be used in Europe, which is based on a meta-analysis written by Hoek et al. in 2013. [Hoek et al., 2013]

The number of attributable deaths is calculated as the difference in the number of deaths between the incidence at the exposure measured over a specific period and that at baseline exposure. [World Health Organization, 2016] The total health risk can be calculated by using zero exposure to PM2.5 as the baseline. There are no thresholds for PM2.5 that has no effect on health. Table 2 summarizes the calculated attributable shares and the relative risks that are calculated from the Finnish data.

Table 2: Attributable Share and Relative Risk

	Min	Max	Mean
Relative risk (RR)	1	1.080	1.027
Attributable share	0	0.074	0.026

5.2 Gini Coefficient

The Gini coefficient is a measure of statistical dispersion and is most often used to represent the income distribution. It measures the inequality between values belonging to a frequency distribution such as the level of income. The coefficient ranges from 0 to 1, if there's no non-negative income. The smaller the coefficient is, the more equally the income is distributed. The Gini coefficient can be defined by the Lorenz curve. The Lorenz curve draws the fraction of the total income of the population on y axis that is the cumulatively earned by bottom x of the population. The 45 degree line therefore represents perfect equality in income distribution. The Gini coefficient can be seen as the ratio of the area that lies between the line of equality and the Lorenz curve (marked A in Figure 6) over the total area under the line of equality (marked BCD).

However, there is one particular weakness with the Gini coefficient. According to Blinder (1980), it's a mechanic measure of inequality, while we are interested in inequality as an indicator of social welfare. If there were two distributions, X and Y, and we found that distribution X assigns less income both to the poorest 20 percent of families and to the richest 20 percent than does distribution Y. Distribution X is more equal at the upper tail (the rich are not quite so rich), but Y is more equal at the lower tail (the poor are not quite so poor). From the viewpoint of the law of diminishing marginal utility, a small increase in income, 50 euro note for example, gives more utility to a poor person struggling for food than it does to a richer person with an abundant lifestyle. However, the Gini coefficient will state, for example, that the Gini ratio for distribution X is 0.36 while that for distribution Y is 0.37. For this reason, distributions with lower Gini coefficients aren't inherently better. [Blinder, 1980]

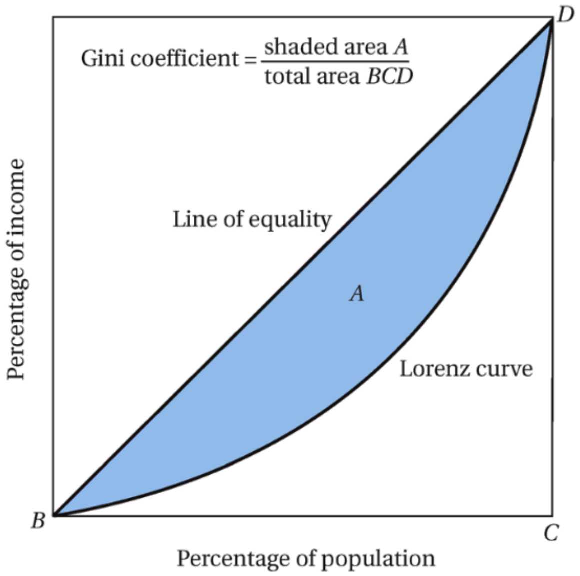

Figure 6: Lorenz Curve [Arndt, 2016]

5.3 Methods

In epidemiology, attributable fraction (or attributable share, a term that will be used in this thesis) among the exposed is the proportion of incidents in the exposed group, that are attributable to the risk factor. In the case of ambient air pollution, the whole population belongs to the exposed group. The attributable risk percent among exposed is used when the fraction is expressed as a percentage. It's calculated as

$$AF_e = (I_e - I_u)/I_e = (RR - 1)/RR \tag{1}$$

where I_e is the incidence in the exposed group, I_u is the incidence in the unexposed group, and RR is the relative risk.

The amount of premature deaths is calculated as

$$(exp(\beta * c) - 1)/exp(\beta * c_r) * \gamma_0 * Pop \tag{2}$$

where β is the concentration-response parameter, c_r is the pollutant (in this case PM2.5) in location r. As the results will be only calculated for year 2015, the time variable t will be left out. γ_0 is the baseline all-cause mortality rate and Pop is the population affected.

The concentration-response parameter is derived from the reported relative risk simply by subtracting value one (1) from it. However, as the CR-parameter is calculated per 10 $\mu g/m^3$, 0.0062 per 1 $\mu g/m^3$will be simply used as it's assumed in epidemiology that the effects grow linearly with the increase in PM2.5 concentration. The increase in PM2.5 exposure is assumed to have an exponential effect to mortality due to exponential models being often observed in biological systems [United States Environmental Protection Agency, 2013].

The concentration-response parameter was selected on the base of the WHO HRAPIE recommendations [World Health Organization, 2015]. The recommended risk coefficient is based on the most recently completed meta-analysis of 13 cohort studies with results by Hoek et al. The studies were conducted in adult populations of North America and Europe. In the case of PM2.5, the concentration-response function used for all-cause mortality implies a relative risk of 1.062 per 10 $\mu g/m^3$. This means that an increase of 10 $\mu g/m^3$ in PM2.5 levels is associated with a 6.2 percent increase in total mortality in the population considered. One downside to this is that it's calculated for ages over 30. There is a short-term exposure RR for all ages (1.0123), but it's not proposed as an alternative to long-term PM2.5 exposure as the long-term exposure already includes the short-term exposure effects. [World Health Organization, 2013]

From previously mentioned function one can derive the attributable share of extant mortality risk that is due to exposure to PM2.5. The same function is also used in the study by Muller et al. (2018) However, they have microdata about the inhabitants' and their household's disposable income while only the average disposable income for adults within a grid-cell is available for this thesis. Instead of calculating the individual attributable share, it will be aggregated to an age cohort level within a grid-cell. The attributable share will be the same for any person belonging to same age cohort in a certain grid.

$$M_{a,r} = \gamma_{0,a} * (1 - 1/exp(\beta * c_r)) \tag{3}$$

where $M_{a,r}$ is the attributable share in age cohort a and location r (grid-cell).

$$\gamma_{0,a\ ,t} = \sum_{i=1} w_i x_i$$

is the weighted average baseline all-cause mortality risk in an age cohort. Here x_i is the all-cause baseline mortality risk for one age x belonging to a certain age cohort (for example, the baseline mortality rate for a 50-year-old, belonging to the age cohort 50-54). One age within an age cohort gets a weight

$$w_i = y_i / \sum_{i=1} y_a$$

where y_i is the amount of individuals of age x and y_a is the total amount of people in one age cohort, where age x belongs (again, the share of 50-year-olds in the age cohort 50-54).

The damage in an age cohort within a grid can be calculated with function

$$D_{a,r} = Pop_a * (VSL * M_{a,r}) \tag{4}$$

where Pop_a is the population within an age cohort.

Then the total damage within a grid is calculated by summing the damages of every age cohort

$$D_r = \sum_{a=1}^{A} D_{a,r} \tag{5}$$

Finally, the total damage is divided by the adult population to get the average cost of damage per grid.

$$D_{avg,r} = D_r / Pop_{adult,r} \tag{6}$$

Adjusted average income for one grid is denoted as I_r^{adj} and defined as the difference between average market income for one grid I_r^m and average damage for one grid $D_{avg,r}$ i.e. $I_r^{adj} = I_r^m$ - $D_{avg,r}$. Adjusted income could be negative for a household if their damages will exceed their incomes, as the study by Muller et al. (2018) found out, but on a grid-level it's not likely.

For that reason, the standard Gini coefficient model will be used with the share of population in one grid to overall population as weights instead of the generalized covariance-based Gini coefficient. When negative adjusted incomes are likely, the standard Gini can't be used as it is restricted by positive values. This standard Gini coefficient will however differ from the official Gini coefficient of Finland due to it being more like a location-based Gini as the income data is available on a grid-level.

This model inherently assumes that most PM2.5 exposure to a person happens in their living area. This means that possible exposure from work place and commuting will be excluded from the study. Another assumption is that the inhabitant has been living in the area for a year or more, due to the fact that CR-parameter is based on a long-term exposure to some level of PM2.5 concentration.

With all the assumptions mentioned above, an example may help the reader to understand the logic behind the model: a 50-year-old person would have been exposed to the grid's PM2.5 level for more than a year. The risk of dying prematurely due to exposure as the person is turning 51 is the attributable share from baseline mortality rate for age cohort 50-54. This risk is given a monetary value by multiplying it with the chosen metric, in this case with VSL (2.65 million euros). This is then done to every person belonging to that age cohort within the same grid.

6 Results

6.1 Main Results

Table 3 reports the estimated Gini coefficients for market income, monetary pollution damage and adjusted income. The VSL average is used as the main metric, while in the sensitivity analysis different metrics demonstrate how it can affect the results. The Gini coefficient for market income is 0.141, which is almost half of the official Gini coefficient of Finland (0.254 in year 2013 and 0.252 in year 2015 [Eurostat, 2020]). This is due to the degree of data precision. The socio-economic data used in this thesis is on a grid level instead of an inhabitant level, which means that one grid has one average income for its inhabitants, wiping out income distribution inside the grid. However, more relevant information is how does the Gini coefficient change from the original when externality from exposure to PM2.5 is included in the measurement.

The Gini coefficient for monetary pollution damage (or externality) is 0.388. The value is higher than the coefficient for market income which suggests that the damages are more concentrated among a subset of grids, at least more than market income. If the grids bearing a greater pollution burden had affluent households, the Gini coefficient for adjusted income might decrease from the original. However, the coefficent for adjusted income increases to 0.155, suggesting that poorer households bear higher monetary pollution damage, at least relative to their income.

The Gini coefficient belongs to sample statistics and therefore should be reported with a standard error. It's often left out due to challenges in computing it. As the Gini coefficient isn't a standard estimator, so the common standard error can't be used. [Giles, 2004] Due to the nature of the Gini coefficient calculated in this thesis, being location-based and thus imprecise, the calculation of standard error is left to be calculated for further studies with more precise microdata of households.

The increase in the Gini coefficient, 0.014 points, would mean that the official Gini coefficient of Finland would increase to 0.266-0.268, assuming that the change would be similar on a different level. This range is close to the Gini coefficients of Denmark and Sweden. Unlike in the Muller et al. (2018)'s study where the Gini coefficient of U.S. increases to the level of South Africa, here the economic significance is quite small. However, the results are likely to be lower bound as the distribution of income within a grid is unknown. Also, the results can be of interest from the viewpoint of economic equality, as they suggest that the monetary cost from exposure does concentrate on average more within areas with poorer population, at least relative to their income.

Table 3: Gini Coefficients

Metric	Market Income	Externality	Market Income - Externality
VSL Average	0.141	0.388	0.155

The grid-level data is divided into income quintiles by their average disposable income, to see how average incomes, damages and pollution levels are distributed. Table 4 compares the estimated shares of average market and adjusted incomes for grids in income quintiles by their average disposable incomes. The results are rounded to three decimal places, which is why the changes in income shares do not match completely. However, from the change we can see that lower income quintiles do indeed experience higher damages relative to their income.

So what could explain the higher damages in the grids with poorer inhabitants? The calculation of damages depends on the PM2.5 concentration level, baseline mortality rate and the monetary value attributed to mortality risk. Because VSL is fixed among grids, the distribution of damages has to be affected by exposure and mortality rates.

Table 4: Income Shares and Average PM2.5 Levels

Income Quintile	Market Income	Adjusted Income	Difference	PM2.5
0-20	0.144	0.139	-0.005	4.801
20-40	0.170	0.166	-0.003	5.029
40-60	0.189	0.188	0.000	5.207
60-80	0.213	0.216	0.002	5.375
80-100	0.285	0.291	0.007	5.618

From the same table we can see that the highest income quintiles are on average exposed to higher level of PM2.5 pollution. Figures 7 and 8 also present the data as a scatter plot, divided into urban and rural municipalities. Urban municipalities include both urban and semi-urban municipalities in which at least 60 per cent live in urban settlements, and in which the population of the largest urban settlement is at least. The remaining municipalities are rural municipalities. [Statistic Finland, 2020b] The logarithmic values of average income and PM2.5 concentration suggest a weak positive correlation between them in both rural and urban municipalities. In Muller et al. (2018)'s study the poorer households were exposed to higher levels of air pollution in United States.

While my results are not completely comparable to theirs, the grid area is quite small which implies that on average richer households are exposed to higher levels of PM2.5 pollution. This result is most likely affected by the situation in Helsinki, where the PM2.5 levels are the highest and the most expensive accommodations, and thus richer population, reside in areas where traffic and traffic congestion is on a high level.

As the pollution levels seem to increase with income on average, something else has to be causing higher damages in the lower income quintiles. The baseline mortality risk is location-invariant as shown in table A. 2, the only variable left to explain the increase in the Gini coefficient is the age of the population in the grids. The baseline mortality risk increases for older age cohorts. Therefore, when VSL is used, the average damages per older age cohorts increase substantially as shown in Figure 9.

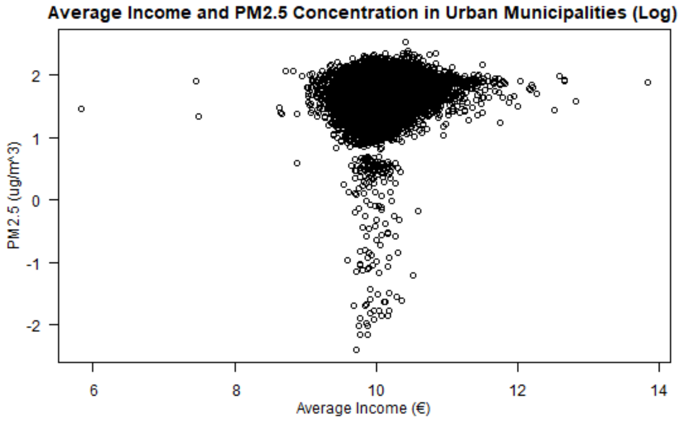

Figure 7: Relation Between Average Income and PM2.5 Levels in Urban Municipalities (Log)

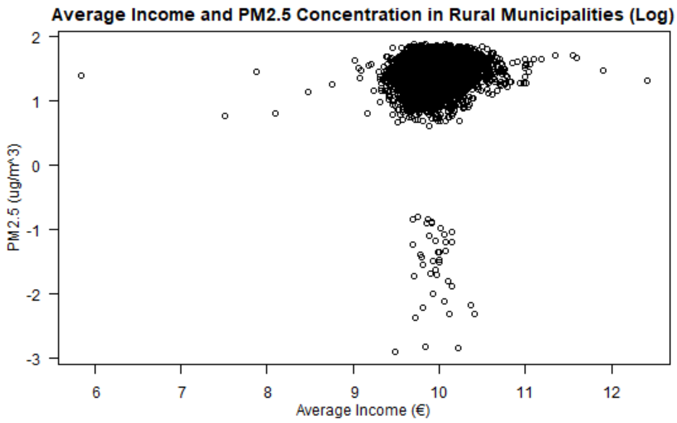

Figure 8: Relation Between Average Income and PM2.5 Levels in Rural Municipalities (Log)

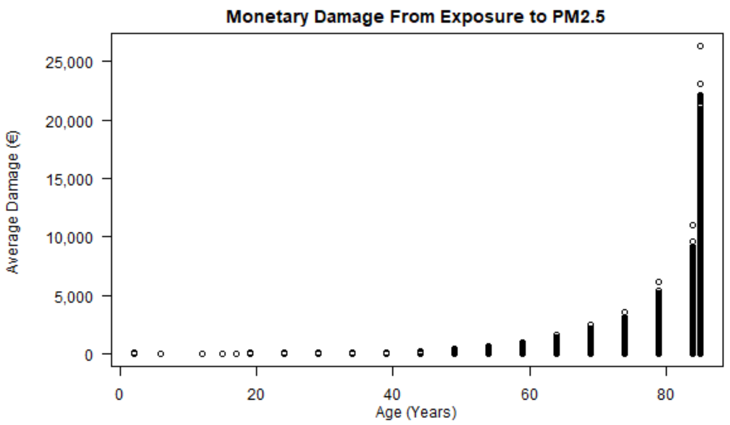

Figure 9: Relation Between Age and Damages

From Figure 10 we can see that the share of older people is higher in the lower quintiles. Therefore, even though the average attributable share to mortality risk is higher for every age cohort in the higher income quintiles, there are less elderly people to be exposed to a higher level of PM2.5 in those grids. This is why the average damage is lower for grids in the higher income quintiles. Table 5 reports the average damages in each quintile.

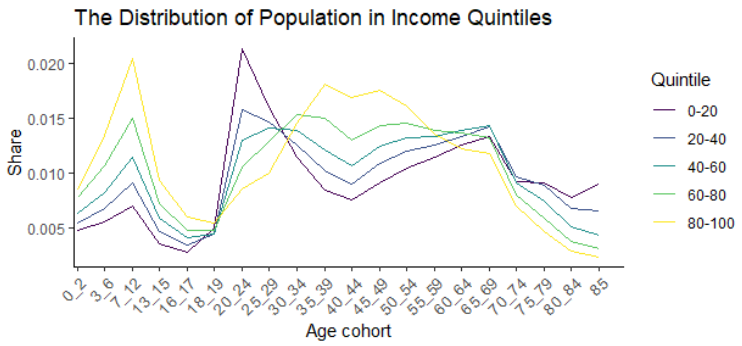

Figure 10: The Distribution of Population in Income Quintiles

Table 5: Average Damages in Income Quintiles

Income Quintile	Min	Max	Mean	Median	SD
0-20	0	15,234	1,346	983	1,313
20-40	0	17,362	1,082	871	911
40-60	0	12,853	883	740	680
60-80	0	11,178	737	610	579
80-100	0	8,290	619	485	524

In supplementary tables A. 3 and A. 4 the summary statistics are presented, separated to the Finnish Regional State Administrative Agencies. The summaries for Damage/Income values are the most informative, as they show how many percent is the average damage relative to average income for the population in a grid belonging to that area. Interestingly enough the maximum share is not the highest in Southern Finland, which has the highest maximum for PM2.5. It is the highest in Western and Inland Finland, where the highest maximum share is 0.934 with the maximum average damage per capita 17,363 euros. On the other hand, the lowest maximum share is in Lapland with the maximum average damage being 6,010 euros. This is mostly likely due to the area having lowest levels of PM2.5 in Finland.

6.2 Reduction Scenario

The European Union has set a National Emission Ceilings Directive that sets national emission reduction commitments for Member States for five air pollutants, including PM2.5, for year 2030. In 2019, the Finnish Ministry of the Environment updated the emission reduction obligations set by the directive which includes reducing PM2.5 with 30 percent till the end of year 2029. [The Finnish Ministry of the Environment, 2019] The relative risk model can be also used to estimate how decreasing PM2.5 levels would affect mortality rates and thus the damages. As mentioned earlier, there is a linear relationship with the concentration-response parameter and the level of pollution, so the saved damages costs can be calculated by decreasing the pollution levels by 30 percent. Unites States Environmental Protection Agency among others has used this method to estimate how climate policies affect the mortality rates [United States Environmental Protection Agency, 2013]. In table 6 the total damages in Finland are summarized. When the VSL metric is used, the savings would be 1.1 billion euros.

Table 6: Total Damages in Finland (€)

Year	Normal Scenario	Reduction Scenario	Difference
2015	3,743,600,000	2,635,500,000	-1,108,100,000

Table 7 shows that the reduction in PM2.5 levels will lead to an decreased Gini coefficient for the adjusted income. From the decreased Gini coefficient and the savings in damages we can separate two channels: income effect and redistribution effect. So, environmental policy aiming to decrease the level PM2.5 can increase the level of economic welfare and change the distribution of said welfare more equal. According to the law of diminishing marginal utility, total welfare would be increased.

Table 7: Gini Coefficients in the Reduction Scenario

Year	Obs.	Market Income	Externality	Market Income - Externality
Original	53,871	0.141	0.388	0.155
Reduced	53,871	0.141	0.388	0.150

Table 8 compares the estimated shares of average market and adjusted incomes for grids in income quintiles by their average disposable incomes when the PM2.5 levels are 30 percent lower. For two of the lowest income quintiles, the relative transfers in the share of income have dropped by 0.1 percentage points.

The highest income quintile experiences the highest effect, as it only gains 0.5 percentage points from the share of total income instead of 0.7. Therefore, the shares become slightly more equal when the pollution levels are decreased.

Table 8: Income Shares With Emission Reduction Scenario

Income Quintile	Market Income	Adjusted Income	Difference
0-20	0.144	0.140	-0.004
20-40	0.170	0.167	-0.002
40-60	0.189	0.188	0.000
60-80	0.213	0.215	0.002
80-100	0.285	0.289	0.005

7 Uncertainty

7.1 Sensitivity Analysis

The point of the sensitivity analysis is to show how different metrics could possibly affect the results. The metrics chosen are the ones mentioned before, the inflation adjusted values from the study NewExt, conducted by The European Environmental Agency. As a reminder, the values for VSL are for one lifetime, while the values for VOLY are per expected life year lost.

As seen in table 9, the value change within one metric only affects the Gini coefficient for adjusted income while the coefficient for the externality stays the same. This is expected, as the value is distributed the same way in both cases. The level of the value affects how much is deducted from the original market income and thus the higher the damage/market income ratio is only for a certain group of population, the higher will be the Gini coefficient for adjusted income.

Table 9: Gini Coefficients With Alternative Metrics

Metric	Value (€)	Market Income	Externality	Market Income - Externality
VSL				
Average	2,650,000	0.141	0.388	0.155
Median	1,300,000	0.141	0.388	0.147
VOLY				
Average	120,000	0.141	0.222	0.147
Median	69,000	0.141	0.222	0.143

When the VOLY metric is used, the Gini coefficient for monetary pollution damage is more equally spread through the population. The coefficients for adjusted income also are lower than when VSL is used. The population is again divided into income quintiles in table 10. Interestingly enough, this time also the middle quintile gains from the share of income. Overall the changes in shares in other quintiles seem to be smaller, which also indicates that the monetary damages are more equally distributed compared to using a VSL metric. However, the differences in the coefficient for adjusted income are 0.004-0.008 points, which is of small magnitude.

The reason why using the VOLY metric results in lower Gini coefficient is due to how it allocates the monetary damages to each age cohort (Figure 11). The younger the person is, the more potential life years are lost. This results in the damage being around one thousand euros for 0-2 -year-old age cohort and 8,000 euros for people over 80.

Table 10: Income Shares With VOLY Average

Income Quintile	Market Income	Adjusted Income	Difference
0-20	0.144	0.141	-0.002
20-40	0.170	0.168	-0.002
40-60	0.189	0.188	0.001
60-80	0.213	0.214	0.001
80-100	0.285	0.289	0.004

With VSL, for the former it's near zero and for the latter 25,000 euros. However, the damage for younger people doesn't still exceed the damage older people are experiencing, even though they lose more years. This is because the mortality rates are anyway a lot smaller for them, which results in quite small damages. The total damage when using VOLY average is 2.9 billion euros, 0.8 billion less than when using VSL average.

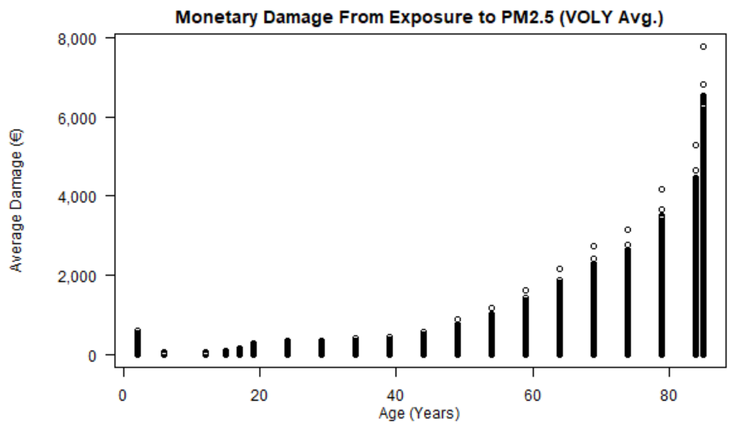

Figure 11: Relation Between Age and Damages

7.2 Robustness

Another possible way to calculate the Gini coefficient for market income without dropping any adult population is to use income information on a postal code level, increasing the included adult population from 3.6 to 4.3 million. However, the coefficient wipes so much variation in disposable income that the Gini coefficient for market income drops to 0.089.

Another concern could be that some of the average incomes could be outliers and affect the Gini coefficient drastically. In table 11,a percentage of randomly chosen incomes are dropped from the calculations. In addition, the result with dropping the highest average income (being an outlier) is presented. The results do not change considerably, as the changes are less than one percentage point.

Table 11: Gini Coefficients With Observation Dropping

Dropped	Obs.	Market Income	Externality	Market Income - Externality
None	53,871	0.141	0.388	0.155
Highest avg. income	53,870	0.140	0.388	0.153
25%	40,403	0.141	0.390	0.155
50%	26,935	0.144	0.384	0.158
75%	13,467	0.138	0.389	0.152

As mentioned before, grids that contain less than 10 people are dropped due to data privacy. How would it affect the results if grids with larger amount of population are dropped? From table 12 we can see that not much. Therefore, the possibility of biased results due to data privacy reasons is low.

Table 12: Gini Coefficients With Grid Dropping

Dropped	Obs.	Market Income	Externality	Market Income - Externality
None	53,871	0.141	0.388	0.155
Less than 11	50,478	0.140	0.387	0.154
Less than 12	47,839	0.140	0.387	0.154
Less than 13	45,579	0.140	0.386	0.154
Less than 14	43,791	0.140	0.385	0.154
Less than 15	42,181	0.140	0.384	0.154
Less than 20	36,364	0.140	0.382	0.154
Less than 25	32,274	0.139	0.380	0.153
Less than 30	28,984	0.139	0.378	0.153
Less than 100	8,665	0.139	0.358	0.152
Less than 500	616	0.133	0.291	0.143

The population data used in earlier mentioned calculations is based on the 2014 age cohort structure, as there is no similar information for year 2013. Therefore, the population and damages are from year 2014, while the average income data is from 2013. If the damages per age cohort are still calculated with the information from year 2014, but otherwise the total amount of adult population per grid from year 2013 is used, how does it affect the distribution of income and damages? The coefficients are presented in table 13 and the income shares in table 14. Again, the effect is small. When using the average VSL, the coefficients decrease by 0.01 percentage points and the income shares do not change at all on three decimal places.

Table 13: Gini coefficients With Adult Population From Year 2013

Metric	Obs.	Market Income	Externality	Market Income - Externality
VSL average	53,871	0.140	0.387	0.154
VSL median	53,871	0.140	0.387	0.147
VOLY average	53,871	0.140	0.221	0.147
VOLY median	53,871	0.140	0.221	0.143

Table 14: Income Shares With Adult Population From Year 2013

Income Quintile	Market Income	Adjusted Income	Difference
0-20	0.144	0.139	-0.005
20-40	0.170	0.166	-0.003
40-60	0.189	0.188	0.000
60-80	0.213	0.216	0.002
80-100	0.285	0.291	0.007

8 Housing Prices and PM2.5 Levels in Helsinki

In this section, the relationship between housing prices and PM2.5 levels will be assessed. It is interesting to see, whether there is any kind of housing segregation between population from different socio-economic backgrounds. This would have significance from the viewpoint of economic equality, if poorer population ended up living in more polluted areas, thus incurring more damage from exposure. The capital of Finland, Helsinki, is of interest as it is an area that is tightly populated and has a lot of variation in it's pollution levels. There are population living near the sea, near highways, in the center and near forest.

There is the possibility that pollution affects housing prices. When all else is equal, we could expect a household to prefer living in a less polluted area, if they take pollution into account when deciding where to live. This should result in higher rental prices and higher property values in cleaner areas. This in turn, would mean that poorer households with income constraints would choose to live in more polluted areas. Therefore a negative correlation between housing prices and PM2.5 levels should be observed. Another nuisance possibly affecting housing prices could be noise pollution. But because noise pollution doesn't really result in premature mortality, it is ruled out of the study.

Rosen (1974) was the first to give an economic interpretation to the correlation between housing prices and air pollution. In his model, a differentiated good can be described by a vector of its characteristics. For a house it could include e.g. number of rooms and air quality. This kind of pricing model is also called hedonic pricing method. In a competitive market, the price function with air pollution characteristic should give an equilibrium differential that allocates individuals across locations and compensates those who face higher pollution levels. Therefore, locations with poor air quality must have lower housing prices in order to attract potential homeowners. [Rosen, 1974][Chay and Greenstone, 2005]

With hedonic pricing model, the marginal price of housing characters is equal to and individual consumer's marginal willingness to pay (MWTP) for that characteristic and an individual supplier's marginal cost of producing it. The MWTP can be used to infer the welfare effects of a marginal change in a characteristic for a given individual. However, consistent estimation of hedonic price schedule (equilibria with consumers' bid functions and suppliers' offer functions) is challenging as there may be unobserved factors that covary with both air pollution and housing prices. Often times areas with higher levels of particulate matter are more urbanised and have higher per capita incomes. [Chay and Greenstone, 2005]

Chay and Greenstone (2005) exploit the Clean Air Act in the U.S. to study how total suspended particulates (TSPs) capitalize into housing values. The legislation imposes strict regulations on polluters in "nonattainment" counties, which are defined by concentrations of TSPs that exceed a federally set ceiling. They find in the study that a nonattainment status is associated with large reductions in TSPs pollution and increases in county-level housing prices. Using the county-level regulations as an instrument, they estimated that a 1 $\mu g/m^3$ reduction in TSPs results in a 0.2–0.4 percent increase in mean housing values. [Chay and Greenstone, 2005]

Bayer et al. (2006) mention that conventional hedonic method for estimating the housing value rely on the assumption that households move freely among locations. But when moving is costly, the variation in housing prices and wages across locations may no longer reflect the value of differences in local amenities. They address the problem by developing an alternative discrete choice approach that models the household location decision directly, and apply it to the case of air quality in U.S. metro areas in 1990 and 2000. Because air pollution is likely to be correlated with unobservable local characteristics such as economic activity, they instrument for air quality using the contribution of distant sources to local pollution. They exclude emissions from local sources which are most likely to be correlated with local conditions. The results show an estimated elasticity of WTP for air quality of 0.34-0.42. This implies that the median household would pay \$149–\$185 (in constant 1982–1984 dollars) for 1 $\mu g/m^3$ reduction in particulate matter. [Bayer et al., 2006]

The average housing price data is based on the prices per square metre of old dwellings in housing companies from Statistics Finland. The prices are on a postal code level and the prices are calculated with 2013-2015 averages. In addition, the PM2.5 concentrations have been aggregated to a postal code averages, based on coordinates. Aggregating the numbers does wipe out variation inside the postal codes, so it is left for further studies to calculate more precise data when grid-level data of housing prices is obtained.

In figure 12 the average housing prices and average pollution are presented in logarithmic values. On a postal code level, one could say that the most expensive houses reside on areas with low to high PM2.5 levels. This could be possible, as there are expensive areas near the sea with low pollution, but also expensive houses at the center with high congestion, increasing PM2.5 levels.

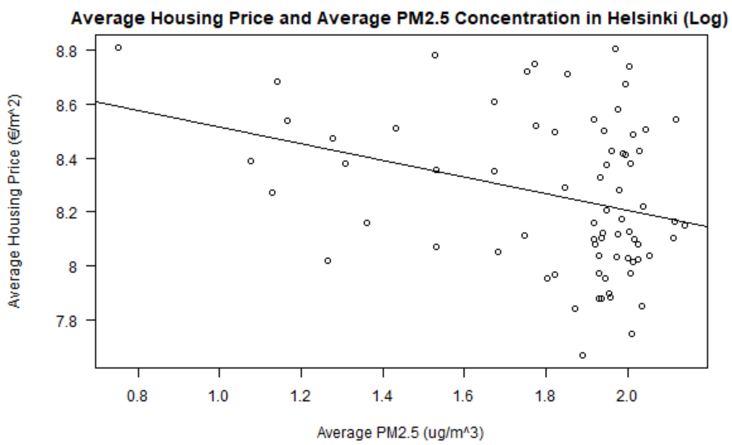

Figure 12: Relation Between Average Housing Price and Average PM2.5 Concentration

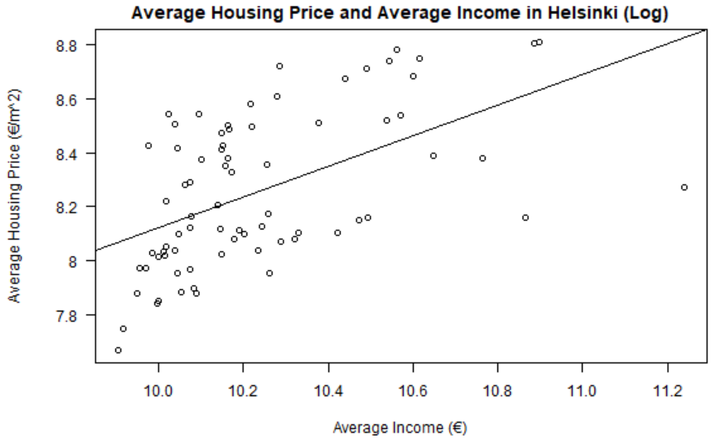

Figure 13: Relation Between Average Housing Price and Average Income

On the other hand, cheaper houses seem to concentrate more on areas with higher levels of PM2.5. Figure 13 describes the relation between average housing prices and average income. As expected, it is likely that people with higher income live in richer areas.

Figures 14 and 15 illustrate the average PM2.5 concentrations (Figure 14) and the average housing prices (Figure 15) in Helsinki. In Figure 15 the areas marked red have the highest prices per square metre. The PM2.5 concentration map shows that the pollution level can be anything from low to high. People with higher income could choose the live in cleaner areas, if the supply of accommodations wasn't constrained. At the same time, the highest average PM2.5 level doesn't exceed the WHO recommendation of 10 $\mu g/m^3$, so it might not be a concern when households are looking for a place to live.

Cheaper accommodations seem to reside on more polluted areas on average. The scatter plot shown earlier also shows a slight negative correlation between housing price and PM2.5 level. The red areas with cheaper accommodation reside nearby highways that naturally increase the pollution levels. Often times these areas are less attractive places to live, which could reflect on housing prices.

From the viewpoint of economic equality, based on the average housing prices and average PM2.5 levels, it is a possibility that income-constrained people can't choose to live in a less polluted area as freely as a richer person would. However, these results are based on averages and the results could change with more precise data.

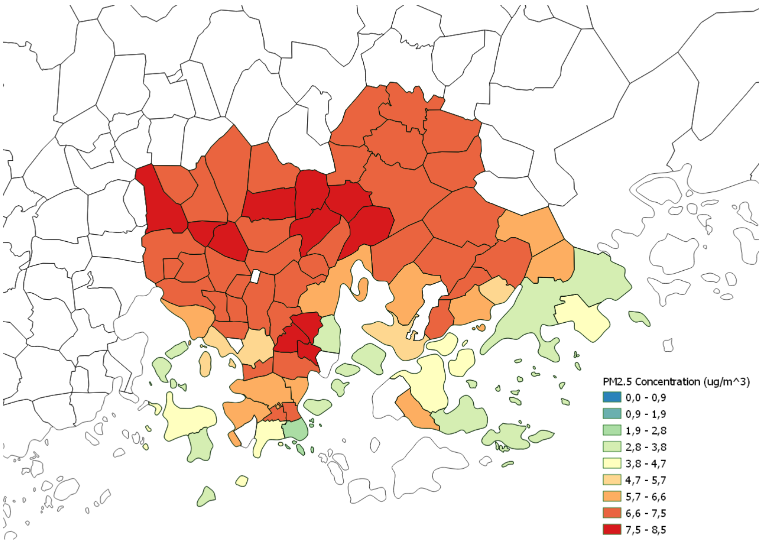

Figure 14: Average PM2.5 Concentrations in Helsinki

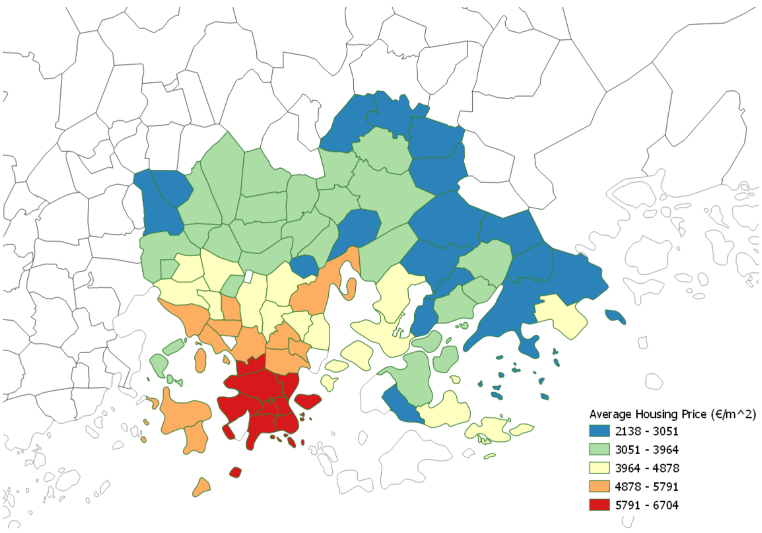

Figure 15: Average Housing Prices in Helsinki

9 Discussion

The results show only a small magnitude, if all, of change in the distribution of adjusted income. However, this might just be due to treating grids as the units of calculations, instead of households. With more precise household income data, the distribution of income would be accurately presented and linked to the pollution data. Now as one grid-cell has one average income, it is assumed that everyone that has turned 18 years old receives the same amount of income.

The results highlight that in general the increase in income inequality is due to lower income quintiles having a larger share of older people who are more vulnerable to air pollution. As the population in Finland is getting older, the amount of premature deaths attributed to exposure is going to increase if the PM2.5 concentration levels won't be lowered.

Concerning the relation of housing prices and PM2.5 levels in Helsinki, cheaper accommodations seem to reside on more polluted areas on average. From the viewpoint of economic equality, based on the average housing prices and average PM2.5 levels, it is a possibility that income-constrained population can't choose to live in a less polluted area as freely as a richer population could, resulting in larger monetary damages from exposure.

What might interest policymakers in Finland is income and redistribution effects the emission reduction policy might have. Not only does the decrease in total damages increase the level of economic welfare in Finland, it also makes the distribution of the damages more equal, as it's not as concentrated on certain group of population. These results could possibly change, however, with more precise data of the household income.

10 Conclusion

In year 2015, air pollution was associated with 2000 premature deaths. Fine particles (PM2.5) were the main contributor (74%) to the disease burden and most of the burden could be attributed to mortality with 1,600 premature deaths and 32,900 years of life lost. The Gini coefficient is calculated for market income, adjusted income and PM2.5 externality. The Value of Statistical Live is used as the main monetary metric as it is fixed, and possible increase in Gini coefficient would be explained by other factors. The calculated Gini coefficient for market income (0.141) is substantially lower than the official Finnish Gini coefficient for year 2013 (0.254). The use of average incomes for a grid wipes out the income distribution inside the grid, underestimating the actual Gini coefficient. However, the most essential thing is the change in Gini coefficient when including an air pollution externality, not the level of the coefficient.

The Gini coefficient increases from 0.140 to 0.155 when adjusted income is used instead of the original market income. The Gini coefficient for externality is 0.388, which means that the damages are concentrated on a subset of grids. By inspecting the population distribution in income quintiles, the results show that the lower quintiles have larger share of older people than the higher ones. Even though the pollution levels are lowest in the lower income quintiles, the inherent vulnerability to air pollution rises the damages high in the lower quintiles.

Damage per market income values show how many percent is the average damage relative to average income for the population in a grid belonging to that area. Interestingly enough the maximum share is not the highest in Southern Finland, which has the highest maximum for PM2.5 levels. The higher the damage/market income ratio is only for a certain group of population, the higher will be the Gini coefficient for adjusted income.

Using VOLY decreases the Gini coefficient for both PM2.5 pollution externality and adjusted income. This is due to the fact that the damages do not concentrate as much on elderly as it does with the VSL metric. However, the sensitivity analysis pointed out that the results do not vary much. less than one percentage point, with different metrics used.

Economic inequality can also show up in housing prices. In Helsinki, cheaper accommodations seem to reside on more polluted areas on average. Based on the average housing prices and average PM2.5 levels, it is a possibility that income-constrained population can't choose to live in a less polluted area as freely as a richer population could, resulting in larger monetary damages from exposure.

The last point to consider is the possible effect of PM2.5 level reduction on mortality risk and it's relation to income distribution. With a reduction of 30 percent in PM2.5 levels, the Gini coefficient for the adjusted income will decrease from 0.155 to 0.150. From the decreased Gini coefficient and the savings in damages we can separate two channels: income effect and redistribution effect. So, environmental policy aiming to decrease the level PM2.5 can increase the level of economic welfare and change the distribution of said welfare more equal. According to the law of diminishing marginal utility, total welfare would be increased.

References

[Arndt, 2016] Arndt, C. (2016). Poverty and Inequality, UNU-WIDER.

[Bayer et al., 2006] Bayer, P., Keohane, N., and Timmins, C. (2006). Migration and Hedonic Valuation: The Case of Air Quality. *Journal of Environmental Economics and Management*.

[Blinder, 1980] Blinder, A. (1980). The Level and Distribution of Economic Well-Being. *The American Economy in Transition, University of Chicago Press for NBER.*

[Chay and Greenstone, 2005] Chay, K. and Greenstone, M. (2005). Does Air Quality Matter? Evidence from the Housing Market. *Journal of Political Economy*, 113(2).

[Cohen et al., 2017] Cohen, A., Brauer, M., Burnett, R., Anderson, H., Frostad, J., Estep, K., Balakrishnan, K., Brunekreef, B., Dandona, L., Dandona, R., Feigin, V., Freedman, G., Hubbell, B., Jobling, A., Kan, H., Knibbs, L., Liu, Y., Martin, R., Morawska, L., Pope III, C., Shin, H., Straif, K., Shaddick, G., Thomas, M., van Dingenen, R., van Donkelaar, A., Vos, T., Murray, C., and Forouzanfar, M. (2017). Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. *The Lancet*, 389:1907–1918.

[Crouse et al., 2012] Crouse, D., Peters, P., van Donkelaar, A., Goldberg, M., Villeneuve, P., Brion, O., Khan, S., Atari, D., Jerrett, M., P., C.A., Brauer, M., Brook, J., Martiv, V.R., Stieb, D., and Burnett, R. (2012). Risk of Nonaccidental and Cardiovascular Mortality in Relation to Long-term Exposure to Low Concentrations of Fine Particulate Matter: A Canadian National-level Cohort Study. *Environmental Health Perspectives*, 120(5).

[Eurostat, 2020] Eurostat (2020). Gini coefficient.

[Giles, 2004] Giles, D. (2004). Calculating a Standard Error for the Gini Coefficient: Some Further Results. *Oxford Bulletin of Economics Statistics*, 66(3).

[Hoek et al., 2013] Hoek, G., Krishnan, R., Beelen, R., Peters, A., Ostro, B., Brunekreef, B., and Kaufman, J. (2013). Long-term Air Pollution Exposure and Cardio-respiratory Mortality: a Review. *Environmental Health*, 12(43).

[Hänninen et al., 2014] Hänninen, O., Knol, A., Lim, T.-A., Conrad, A., Rappolder, M., Carrer, P., Fanetti, A.-C., Kim, R., Buekers, J., Torfs, R., Iavarone, I., Classen, T., Hornberg, C., and Mekel, O. (2014). Environmental Burden of Disease in Europe: Assessing Nine Risk Factors in Six Countries. *Environmental Health Perspectives*, 122(5).

[Karvosenoja, 2008] Karvosenoja, N. (2008). Emission Scenario Model for Regional Air Pollution. *Monographs of the Boreal Environment Research*, (32).

[Laden et al., 2006] Laden, F., Schwartz, J., Speizer, F., and Dockery, D. (2006). Reduction in Fine Particulate Air Pollution and Mortality: Extended Follow-up of The Harvard Six Cities Study. *American Journal of Respiratory and Critical Care Medicine*, 173(6):667–672.

[Lehtomäki et al., 2018] Lehtomäki, H., Korhonen, A., Asikainen, A., Karvosenoja, N., Kupiainen, K., Paunu, V.-V., Savolahti, M., Sofiev, M., Palamarchuk, Y., Karppinen, A., Kukkonen, J., and Hänninen, O. (2018). Health Impacts of Ambient Air Pollution in Finland.

[Muller et al., 2018] Muller, N., Matthews, P., and Wiltshire-Gordon, V. (2018). The Distribution of Income Is Worse Than You Think: Including Pollution Impacts Into Measures of Income Inequality. *PLoS ONE*, 13(3).

[National Academy of Sciences, National Research Council, 1999] National Academy of Sciences, National Research Council (1999). Nature's Numbers: Expanding the U.S. National Economic Accounts to Include the Environment. *National Academy Press*.

[Nordhaus, 2006] Nordhaus, N. (2006). Principles of National Accounting for Non-Market Accounts. *A New Architecture for the U.S. National Accounts*.

[Nordhaus and Tobin, 1972] Nordhaus, W. and Tobin, J. (1972). Is Growth Obsolete? Economic Research: Retrospect and Prospect. *Economic Research: Retrospect and Prospect*, 5.

[Rosen, 1974] Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy*, 82(1):34–55.

[Savolahti et al., 2018] Savolahti, M., Kangas, L., Karppinen, A., Karvosenoja, N., Kukkonen, J., Lanki, T., Nurmi, V., Palamarchuk, Y., Paunu, V.-V., Sofiev, M., and Tiittanen, P. (2018). Ilmansaasteiden haittakustannusmalli Suomelle (IHKU). *Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisusarja 26/2018*.

[Soimakallio et al., 2017] Soimakallio, S., Hildén, M., Lanki, T., Eskelinen, H., Karvosenoja, N., Kuusipalo, H., Lepistö, A., Mattila, T., Mela, H., Nissine, A., Ristimäki, M., Rehunen, A., Repo, A., Salonen, R., Savolahti, M., Seppälä, J., Tiittanen, P., and Virtanen, S. (2017). Energia- ja ilmastostrategian ja keskipitkän aikavälin ilmastopolitiikan suunnitelman ympäristövaikutusten arviointi. *Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisusarja*, 59.

[Statistic Finland, 2020a] Statistic Finland (2020a). Disposable Money Income.

[Statistic Finland, 2020b] Statistic Finland (2020b). Grid Database 2015. Content Description.

[Statistic Finland, 2020c] Statistic Finland (2020c). Probability of Death.

[Statistic Finland, 2020d] Statistic Finland (2020d). Property Income.

[The Finnish Ministry of the Environment, 2019] The Finnish Ministry of the Environment (2019). Kansallinen ilmansuojeluohjelma 2030.

[United States Environmental Protection Agency, 2013] United States Environmental Protection Agency (2013). Regulatory Impact Analysis for the Final Revisions to the National Ambient Air Quality Standards for Particulate Matter.

[Vähänen, 2015] Vähänen, M. (2015). Tupakoinnin yhteiskunnalliset kustannukset ja niiden arviointimenetelmät.

[World Health Organization, 1958] World Health Organization (1958). Expert Committee on Environmental Sanitation & World Health Organization. Air Pollution: Fifth Report of the Expert Committee on Environmental Sanitation.

[World Health Organization, 1987] World Health Organization (1987). Air Quality Guidelines for Europe. Copenhagen: WHO Regional Office for Europe.

[World Health Organization, 2000] World Health Organization (2000). Air Quality Guidelines for Europe, 2nd ed.; WHO Regional Publications, European Series.

[World Health Organization, 2006] World Health Organization (2006). WHO Air Quality Guidelines for Particulate Matter, Ozone, Nitrogen Dioxide and Sulfur Dioxide: Global Update 2005: Summary of Risk Assessment.

[World Health Organization, 2013] World Health Organization (2013). Health Risks of Air Pollution in Europe – HRAPIE project.

[World Health Organization, 2015] World Health Organization (2015). Quantifying the Health Impacts of Ambient Air Pollutants: Recommendations of a WHO/Europe Project.

[World Health Organization, 2016] World Health Organization (2016). Health Risk Assessment of Air Pollution - General Principles. *Copenhagen: WHO Regional Office for Europe.*

[World Health Organization, 2018] World Health Organization (2018). Ambient (Outdoor) Air Quality and Health.

[World Health Organization, 2020] World Health Organization (2020). Metrics: Disability-Adjusted Life Year (DALY).

A Appendix

Table A. 1: Population Structure, year 2014

	Min	Max	Mean	Total	Obs.
Inhabitants, total	1	2,745	517	5,407,760	327,805
0-2 years	0	105	16	176,086	327,805
3-6 years	0	114	23	244,777	327,805
7-12 years	0	136	33	353,615	327,805
13-15 years	0	78	16	173,828	327,805
16-17 years	0	105	11	117,922	327,805
18-19 years	0	56	12	126,412	327,805
20-24 years	0	702	32	336,417	327,805
25-29 years	0	575	31	333,025	327,805
30-34 years	0	497	33	348,873	327,805
35-39 years	0	258	32	338,840	327,805
40-44 years	0	162	29	309,981	327,805
45-49 years	0	150	33	354,480	327,805
50-54 years	0	162	35	371,017	327,805
55-59 years	0	119	35	366,973	327,805
60-64 years	0	103	35	371,914	327,805
65-69 years	0	138	35	373,961	327,805
70-74 years	0	102	22	237,583	327,805
75-79 years	0	80	18	197,439	327,805
80-84 years	0	91	13	142,457	327,805
85 years or over	0	215	12	132,160	327,805

Table A. 2: Baseline Mortality Risk by Age Cohorts, year 2015

	Baseline Mortality Rate (per mil)
0-2 years	0.642
3-6 years	0.060
7-12 years	0.078
13-15 years	0.120
16-17 years	0.195
18-19 years	0.392
20-24 years	0.494
25-29 years	0.541
30-34 years	0.675
35-39 years	0.824
40-44 years	1.192
45-49 years	2.088
50-54 years	3.182
55-59 years	5.152
60-64 years	8.105
65-69 years	12.509
70-74 years	18.085
75-79 years	31.448
80-84 years	55.743
85 years or over	133.285

Table A. 3: Summary Statistics by Region

Region	Average Income	PM2.5	Per Capita Damage	Damage /Income
Åland				
Mean	26,665	3.82	684	0.029
Median	25,552	5.06	445	0.017
SD	11,386	2.30	934	0.046
Min.	13,172	0	0	0
Max.	202,806	5.90	8,558	0.422
Obs.	339	339	339	339
Southern Finland				
Mean	26,356	5.89	912	0.040
Median	24,468	5.81	722	0.029
SD	14,241	1.21	816	0.0435
Min.	8,301	0	0	0
Max.	1,017,969	12.49	12,853	0.662
Obs.	18,154	18,154	18,154	18,154
Eastern Finland				
Mean	21,768	4.53	932	0.048
Median	20,829	4.39	711	0.034
SD	7,029	0.68	851	0.051
Min.	5,624	3.11	37	0
Max.	274,982	7.64	9,037	0.503
Obs.	5,669	5,669	5,669	5,669
Western and Inland Finland				
Mean	22,958	5.20	948	0.046
Median	22,090	5.01	726	0.032
SD	6,083	1.19	898	0.051
Min.	345	0	0	0
Max.	245,545	10.04	17,363	0.934
Obs.	13,841	13,841	13,841	13,841

Table A. 4: Summary Statistics by Region

Region	Average Income	PM2.5	Per Capita Damage	Damage /Income
Lapland				
Mean	22,135	3.27	568	0.028
Median	21,624	3.31	432	0.020
SD	5,004	0.98	549	0.030
Min.	340	0	0	0
Max.	77,006	6.31	6,010	0.318
Obs.	2,529	2,529	2,529	2,529
Southwestern Finland				
Mean	23,795	5.79	1,076	0.051
Median	22,663	5.80	830	0.036
SD	11,357	1.19	1,032	0.058
Min.	1,713	0	0	0
Max.	603,017	9.33	12,855	0.747
Obs.	7,674	7,674	7,674	7,674
Northern Finland				
Mean	22,650	4.15	679	0.033
Median	21,837	4.08	511	0.023
SD	5,908	0.80	653	0.037
Min.	1,777	0	0	0
Max.	148,286	8.69	8,031	0.467
Obs.	5,665	5,665	5,665	5,665

Table A. 5: Adult Population

	2014	2013
Included	3,558,553	3,543,896
Total	4,341,532	4,321,588

Table A. 6: Adult population by Region

Region	Population
Åland	13,481
Southern Finland	1,627,295
Eastern Finland	326,839
Western and Inland Finland	751,125
Lapland	106,232
Southwestern Finland	450,029
Northern Finland	283,552